eMSQOL-29: Prospective validation of the abbreviated, electronic version of MSQOL-54

Full title:

eMSQOL-29: Prospective validation of the abbreviated, electronic version of the MSQOL-54

Authors:

Rosato R[1], Testa S[1], Bertolotto A[2], Scavelli F[2], Giovannetti AM[3,4], Confalonieri P[4], Patti F[5], Chisari CG[5], Lugaresi A[6], Pietrolongo E[6], Grasso MG[7], Rossi I[7], Toscano A[1], Loera B[1], Giordano A[1,3], Solari A[3]

1. Department of Psychology, University of Turin, Turin, Italy
2. Regional Referral Multiple Sclerosis Centre (CReSM), University Hospital San Luigi Gonzaga, Orbassano, Italy
3. Unit of Neuroepidemiology, Foundation IRCCS Neurological Institute C. Besta, Milan, Italy
4. Department of Neuroimmunology and Neuromuscular Diseases, Foundation IRCCS Neurological Institute C. Besta, Milan, Italy
5. MS Centre, Neurology Clinic, University Hospital Policlinico Vittorio Emanuele, Catania, Italy
6. Department of Neuroscience, Imaging and Clinical Sciences, G. d'Annunzio University of Chieti-Pescara, Chieti, Italy
7. Multiple Sclerosis Unit, IRCCS S. Lucia Foundation, Rome, Italy

Corresponding author:

Alessandra Solari, MD

Unit of Neuroepidemiology

Foundation IRCCS Neurological Institute C. Besta

Via Celoria 11,

20133 Milan, Italy

E-mail: alessandra.solari@istituto-besta.it

Tel. +390223944016; Fax +390223944056

Keywords: Health-related quality of life; Multiple sclerosis; MSQOL-54; electronic migration, patient reported outcomes, psychometric assessment

Running title: Prospective validation of the abbreviated, electronic version of the MSQOL-54

ABSTRACT

Background: We recently devised a shortened version of the 54-item Multiple Sclerosis Quality of Life (MSQOL-54) in paper (MSQOL-29, consisting of 25 items forming 7 subscales, and 4 single items; one filter question for 3 'sexual function' items) and electronic format (eMSQOL-29).

Objectives: To prospectively assess eMSQOL-29 psychometric properties, acceptability/equivalence vs. MSQOL-29.

Methods: MS patients (n=623; EDSS range 0.0–9.0) completed eMSQOL-29, Hospital Anxiety and Depression Scale, Functional Assessment of MS (FAMS), European Quality of life Five Dimensions-3L, and received EDSS and Symbol Digit Modality Test (SDMT). Equivalence vs. MSQOL-29 was assessed in 242 patients (randomized cross-over design).

Results: 'Sexual function' items were filtered out by 273 patients (47%). No multi-item scale had floor effect, while 5 had ceiling effect. Cronbach's alpha range was 0.88–0.90. Confirmatory factor analysis showed good overall fit, and the two-factor solution for composite scores was confirmed. Concurrent validity was sub-optimal for 'cognitive function' (vs. SDMT, r=0.25) and 'social function' (vs. FAMS social function, r=0.38). eMSQOL-29 equivalence was confirmed, and its acceptability was good.

Conclusions: eMSQOL-29 showed good internal consistency, factor structure, no floor effect, while most subscales had some ceiling effect. Concurrent validity was sub-optimal for two subscales. Equivalence and acceptability were good.

INTRODUCTION

Multiple sclerosis (MS) is a chronic disease of the central nervous system characterized by a multiplicity of symptoms and signs, and a variable impact on physical, psychological and social functioning. Health-related quality of life (HRQOL) and other patient-reported outcomes (PROs) are key for assessing the disease from the patient's perspective [1-3]. However, it is recommended that PRO questionnaires should be as brief as possible in order to minimize respondent burden [4, 5]. For MS patients in particular, presence of fatigue and impairment of sustained attention are common symptoms, which can limit patient's ability to complete long PRO instruments [6]. From 1995 a number of MS-specific HRQOL instruments have been produced [6-9]. With 51 translations available [https://eprovide.mapi-trust.org/instruments/multiple-sclerosis-quality-of-life-54], the 54-item MS Quality of Life inventory (MSQOL-54) is the most-used MS-specific questionnaire [10]. In a previous study phase, we shortened the MSQOL-54 using a combination of psychometric analyses (factor analysis and Rasch modelling) applied on a dataset of 635 MS patients who completed the MSQOL-54, and input from MS/HRQOL professionals and MS patients [11]. The resulting MSQOL-29 consists of 7 multi-item and 4 single-item subscales, used to form two composite scores, consistent with the theoretical construct used to develop the original instrument [12]. The MSQOL-29 requires approximately 10 minutes to complete, corresponding to half of the MSQOL-54 completion time [13]. Nonetheless, the high correlation of MSQOL-29 subscale and composite scores with those of MSQOL-54 suggests that eliminating items and subscales did not substantially change the HRQOL dimensions found for the original instrument. The electronic version provides an alternative administration mode, and can be integrated within electronic health records and disease registries. Another advantage of electronic mode of administration is the score calculation routine which, by reducing computation burden and errors, should facilitate the use of HRQOL measures in clinical practice, and ultimately the patient-

provider communication and shared decision-making. However, to have confidence in the validity of data collected using different administration modes, and to allow pooled analyses when different versions are used within and between studies, between-mode equivalence and acceptability must be formally evaluated [14-16].

Here are presented the results of a prospective study aimed to assess MSQOL-29 acceptability, validity, and reliability, and to confirm its factorial structure in an independent population. In addition, we assessed the equivalence of the electronic version to the paper version.

METHODS

Participants

We used a cross-sectional design with consecutive sampling, and a nested longitudinal equivalence study. Participants were included with a clinical diagnosis of MS[17] and were at least 18 years of age, able to read and understand Italian, and to provide written informed consent. Recruitment occurred at 5 centres across Italy (Foundation IRCCS Neurological Institute C. Besta, Milan; University Hospital San Luigi Gonzaga, Orbassano, Turin; IRCCS S. Lucia Foundation, Rome; G. d'Annunzio University of Chieti-Pescara, Chieti; University Hospital Policlinico Vittorio Emanuele, Catania). The study was approved by the ethics committee of the 5 participating centres.

PRO measures

The MSQOL-29 consists of 7 multi-item subscales: ‘physical function’ (6 items); ‘sexual function’ (4 items); ‘bodily pain’, ‘emotional wellbeing’, ‘energy’, ‘cognitive function’, and ‘health distress’ (3 items); and four single-item subscales (‘social function’, ‘health perceptions’, ‘overall quality of life’, and ‘change in health’) which form two composite scores (Physical Health Composite, PHC;

Mental Health Composite, MHC) [11]. A filter question (*'During the past 4 weeks, have you had an active sexual life?'*) is present after the first 'sexual function' item ('*How much of a problem was lack of sexual interest for you during the past 4 weeks?')*. If the reply is "no", the other 3 'sexual function' items are not shown (eMSQOL-29) or skipped (paper format). An integrated scoring routine is available for the eMSQOL-29, to ease score calculation and interpretation.

The FAMS consists of a generic core HRQOL measure (the Functional Assessment of Cancer Therapy-General scale), supplemented with MS specific items [18]. The 59 items of the FAMS are divided into 6 subscales: 'mobility', 'symptoms', 'emotional well-being', 'general contentment', 'thinking/fatigue' and 'family/social well-being'. Only items 1–44 are included in the total score (0–4 score range for each), which can range from 0 to 176 (best HRQOL).

The European Quality of life Five Dimensions-3L (EQ-5D-3L) is a standardized instrument providing a generic measure of health, via a descriptive system addressing 5 health state dimensions ('mobility', 'self-care', 'usual activities', 'pain/discomfort', and 'anxiety/depression'), each scored on a three-level scale. A visual analogue scale (VAS) is also included to measure overall health. By combining every characteristic level of the descriptive system, a total of 243 profiles can be obtained, within the range 11111 to 33333 (worst health-state). The profiles can be assigned an index score (utility value, available for several countries included Italy) which represents preferences for that profile derived from a general population survey [19, 20]. The index has a maximum value of 1 (full health), an anchor of 0 for a state equivalent to being dead, and values <0 for states regarded as worse than being dead.

The HADS is a screening tool comprising 14 multiple-choice items (0–3 score range for each), 7 probing symptoms of anxiety and 7 of depression. HADS Anxiety and Depression scores are obtained, which can range from 0 to 21 (most severe symptoms) [21].

Procedure

All participants signed the informed consent form, then completed the eMSQOL-29[11], a section on questionnaire acceptability (see below), and (in citation order), Italian versions of the following inventories: the HADS[22], the FAMS[23], and the EQ-5D-3L. All the questionnaires were completed by the patients on the PC. The MS clinician administered the Expanded Disability Status Scale (EDSS),[24] the symbol digit modality test (SDMT)[25], and recorded patient general and clinical information. eMSQOL-29 acceptability and usability were assessed by rating (0-10 VAS) the suitability of the MSQOL-29 as a MS-specific QOL inventory; and the following four characteristics: being taxing, boring, difficult to read, and difficult to complete. Each characteristic was rated on a 0-3 Likert scale, followed by a multiple-choice question exploring the main reasons for difficulty.

A minimum of 200 participants from Milan and Orbassano centres completed both paper and eMSQOL-29 in random order, and in two-week interval. Re-test was performed at patient home (except for patients preferring to complete at the MS centre). At the end of the visit, participants assigned to the electronic-paper order were given a closed envelope, to be opened after two weeks, containing the MSQOL-29 questionnaire and a pre-addressed, return-paid envelope. A few days before re-test date, participants assigned to the paper-electronic order received an email with the link to the website containing the eMSQOL-29.

After re-testing, eMSQOL-29 acceptability was assessed as reported above; in addition one multiple-choice item assessed the preferred version (electronic, paper, no preference).

One week after the expected completion/return date, patients who did not complete/return the questionnaire received a reminder (email or phone call).

Statistical analysis

Internal reliability was assessed using Cronbach's alpha (benchmark value >0.70). For each MSQOL-29 subscale, we determined presence of floor effect (>10% of the patients scoring minimum) and ceiling effect (>10% of the patients scoring maximum). Concurrent validity was evaluated by correlating MSQOL-29 subscale scores and pertinent subscale/scale scores of FAMS, EQ-5D-3L, and HADS (Pearson's r). We conducted a confirmatory factor analysis (CFA) to test the two-factor structure proposed by Vickrey et al. for the MSQOL-54,[10] using the following cutoff criteria for the goodness of fit indices: Root Mean Square Error of Approximation (RMSEA) <0.05, Comparative Fit Index (CFI) ≥0.95, Standardized Root Mean square residual (SRMR) ≤0.08. Criterion validity of MSQOL-29 subscales/scales was assessed using the EDSS (two groups, cutoff point 4.5) and the HADS (two groups, cutoff point 8.0 for both anxiety and depression) as criterion variables. Independent sample t-test and Cohen's *d* were used for these comparisons. A *d* of 0.50 corresponds to a moderate effect size, and a *d* of 0.80 to a large effect size.

For each of the 11 MSQOL-29 subscale scores, equivalence of paper and electronic version was assessed using a randomized cross-over design, which allows for testing of version (paper, electronic), order (test, retest) effects, and their interactions (sequence). The required sample size (185 patients) was calculated under the following assumptions: intraclass correlation coefficient (ICC) [26] between versions of 0.80 (vs. a reference value of 0.70), 80% power of detecting this difference, 5% alpha error, two-tailed test, and 20% missing items plus drop-outs [27]. In addition to the ICC, we assessed the effect size statistic (benchmark value ≤0.20),[28] and linear regression model for repeated measures design, with an exchangeable correlation structure of the errors. The model included the following independent variables: MSQOL-29 version, order, sequence, centre (Milan, Orbassano), sex, age, EDSS (≤2.5, >2.5) and disease course (relapsing-remitting, primary or secondary progressive).

All the analyses were performed with SAS version 9.4 and LISREL version 8.72. Significance level was set at 0.05.

RESULTS

Participants' characteristics

Between September 2015 and May 2016, 623 adult MS patients were assessed (Table 1). Of these, 424 (68%) were women, the mean age was 44 years (range 20–78). Most had relapsing-remitting MS (82%), median EDSS score was 2.5 (range 0.0–9.0). General and clinical characteristics of the patients were similar across the centres, except for a lower proportion of relapsing-remitting MS (53%) and a higher EDSS score (median 6.0) in Rome (Table 1), reflecting the case-mix of a rehabilitation centre. Forty-one patients (7%) were excluded from the analyses because of incomplete data. A total of 233 MS patients (Milan n=92, Orbassano n=141) participated in the longitudinal equivalence study.

Psychometric properties of the eMSQOL-29

Missing replies for the eMSQOL-29 ranged from 0.2% to 5.7% (item 29), and 3 of the 4 'sexual function' items were filtered out by 273 patients (47%). After confirmation of the invariance of the instrument in the two subsamples (online supplementary appendix A), the 'sexual function' subscale score was computed using all the available information (i.e. one item or 4 items, depending on response to the filter question). As expected, respondents who skipped the sexual items were on average older and with a more severe disease (Table 2).

There was no floor effect for any multi-item scale, while a ceiling effect was present for 'physical function', 'sexual function', 'cognitive function', 'bodily pain' and 'health distress'. Cronbach's alpha ranged from 0.88 to 0.90 (Table 3).

The CFA of the two-factor solution indicated reasonably adequate fit (RMSEA 0.055; CFI 0.99; SRMR 0.035) and outperformed the one-factor solution (Satorra-Bentler scaled chi-square test [1] 14.25; p<0.0001). MHC and PHC scores were derived by analogy with those of the MSQOL-54 as weighted sums of the corresponding subscales (Table 4).

Correlations between eMSQOL-29 and FAMS subscales addressing similar domains ranged from 0.62 to 0.86, except for 'social function' (0.38; p<0.001). Correlation between 'emotional wellbeing' and HADS subscales was -0.59 for Anxiety (p<0.0001), and -0.56 for Depression (p<0.0001). Correlation between 'cognitive function' and SDMT was poor (0.25; p<0.0001). Finally, 'overall quality of life' correlated highly with the EQ-5D VAS (0.66; p<0.0001), than with the EQ-5D-3L index (0.48; p<0.0001; table 5).

Radar plot of MSQOL-29 scale/subscale scores by EDSS score (cut-off value 4.5) is reported in Figure 1. A large effect size (Cohen's d > 0.80) was found for 'physical function', 'overall quality of life', 'energy', 'social function', and for both composite scales. Radar plots of MSQOL-29 scale/subscale scores by HADS Anxiety and Depression scores (cut-off value 8.0) are reported in Figure 2. For HADS Anxiety, a large effect size was found for 'health distress' and for the mental health composite scale; for HADS Depression, it was found for 'health distress', 'emotional wellbeing', 'energy', 'cognitive function' , 'overall quality of life', 'physical function', and for both composite scales.

eMSQOL-29 acceptability

The overall mean rating of the eMSQOL-29 acceptability was 6.9 (SD 1.8), with about 75% of patients (N=426/564) considering it as suitable (VAS score ≥6). There were no significant differences in the evaluation according to the patient clinical and demographic characteristics (i.e. age, education, EDSS, and centre; online supplementary figure). Of 581 patients, 16 (2.8%)

considered the questionnaire as being taxing, 33/579 (5.7%) as boring. None of the patients found the questionnaire difficult to read or complete.

Equivalence study

Of the 233 MS patients enrolled, 220 (94%) completed both questionnaire versions, and 13 did not return the paper MSQOL-29. Missing replies for the paper version ranged from 1 (0.5%) to 5 (2%, item 1); for the eMSQOL-29, were from 1 (0.45%) to 12 (5%, item 29).

Eight of the 11 subscales had ICC values ≥0.70 (table 6), with higher values for multi-item subscales (median ICC 0.87, range 0.84–0.95) compared to single-item subscales (median ICC 0.63, range 0.52–0.76). All the effect sizes were <0.20 (range 0.003–0.12; table 6).

Effect sizes and ICCs for the individual 29 items are reported in the online supplementary table 1.

MSQOL-29 version, order and sequence of administration did not affect the subscale scores in the linear mixed model regression analyses (table 7). Finally, 87 (40%) did not express any preference about the version, 2 missed the answer, while the remaining 131 were equally divided between those who preferred the paper (n=65) and eMSQOL-29 (n=66).

DISCUSSION

We recently devised an abbreviated version of the MSQOL-54, also available in electronic version (eMSQOL-29) using a combination of factor analysis and Rasch modelling. The conceptualization of the MSQOL-29 is similar to the parent HRQOL scales, the MSQOL-54 and the SF-36: 'physical function', 'cognitive function', 'sexual function' and 'social function' subscales aim to capture the behavioural consequences of the health problem (here MS); 'emotional wellbeing', 'bodily pain', 'energy', 'health distress', 'health perceptions' and 'change in health' aim to reflect more

subjective components of health; and ‘overall quality of life’ intends to capture overall satisfaction with life and well-being [31].

In the present study we confirmed eMSQOL-29 reliability, validity and factorial structure in a prospective, independent MS patient sample. In addition, we demonstrated eMSQOL-29 acceptability, and equivalence with the paper version.

An unexpected finding was the high proportion of patients (47%) who filtered out the ‘sexual function’ items, which outnumbered missing items of the parental questionnaire, where the filter question is not present [11]. Acknowledging the worth of addressing sexuality in MS,[10, 29, 30] exploring sexual functioning over a 4-week period can be challenging, especially for the older and more disabled patients (Table 2). In these patients, using a single item (item 24) can be worthwhile. Importantly, by demonstrating the measurement invariance of the ‘sexual function’ subscale score obtained from item 24 vs. the full set of items (items 24-27), we supported the use of this scale in the proposed adaptive version. However, an MS-specific instrument, such as the Multiple Sclerosis Intimacy and Sexuality Questionnaire-15 can be added to the eMSQOL-29 to explore in more detail sexual functioning [32] in MS patients.

Concerning missing items, the last eMSQOL-29 item (item 29, ‘overall quality of life’) had 5% of missing replies (vs. no missing replies in the paper version). Item 29 is the only item in which response is obtained by moving the mouse along a VAS, which can be demanding to patients with tremor or impaired arm function.

Concurrent validity, assessed using various instruments, was acceptable, the only two exceptions being the suboptimal correlations between ‘social function’ and FAMS ‘family/social wellbeing’, and between ‘cognitive function’ and SDMT. Concerning the social function domain, this can originate from differences in item contents between the questionnaires: MSQOL-29 ‘social function’ referring to social activities with family, friends, and social relations; by contrast, FAMS

‘family/social wellbeing’ focuses on social activities within the family. To better appraise the reason for the low correlation between MSQOL-29 ‘cognitive function’ and SDMT (r=0.25; p<0.0001), we assessed the correlation between FAMS ‘thinking/fatigue’ and SDMT, which was also low (r=0.24; p<0.0001). This finding indicates that the most used screening measure of MS cognitive compromise little correlated with patient self-perceived impact of cognitive problems on daily functioning. Poor correlations between subjective and objective measures of cognitive impairment have been reported also elsewhere [33,34]. As expected, the EDSS differentiated well MSQOL-29 subscale scores, chiefly the ‘physical function’ (Figure 1), and depressive symptoms differentiated most subscales, and both composite scales (Figure 2).

The equivalence between the eMSQOL-29 and the paper version was good, with ICC values >0.70 in eight of the 11 subscales, and negligible effect sizes, supporting comparison and pooled analysis of data obtained from the two modes of administration [35].

Study limitations

Power calculation was made only for the equivalence study. However, the size of our clinical validation sample was quite large, and the sample well-varied in terms of disease severity, duration, and socio-demographic characteristics.

Another limitation is that usability and pilot testing were not enough meticulous. Thanks to the study findings, refinements will be made to the eMSQOL-29 to improve its usability by patients with impaired arm function. Specifically, a picker wheel widget will be used to rate the VAS; after rating, an automated feedback message will appear, reporting the selected VAS score; and an alerting message (in case of missing rating) to prevent inadvertent missing replies. These refinements are minor modifications that do not alter the content, meaning, or interpretation at

the item or scale level, and will not require any formal validation besides cognitive debriefing and usability testing [14, 15].

To conclude, the eMSQOL-29 (with integrated scoring routine) has good psychometric properties and is equivalent to the paper version. The good level of agreement between electronic and paper versions should be reassuring to investigators, authorities and sponsors using electronic devices to collect PRO data, having implications for the use of electronic measures [15]. Both versions can be easily used at the international level in clinical practice and research. Work needs to be carried out within a longitudinal design to assess instrument's responsiveness - the ability to detect change over time in the construct to be measured [36], and response shift - where the meaning of scores changes over time as patients adapt to their illness [37]. For this reason, it is premature to recommend use of the MSQOL-29 as an outcome measure in clinical trials. However, its conceptual and methodological strengths, combined with the good psychometric properties here described suggest that eMSQOL-29 may have a place among the MS-specific HRQOL instruments. Further, questionnaire brevity limits compilation burden, real-time scoring prevents scoring burden and errors, and supports use during the consultation to promote shared decision making and patient-centred care [38].

ACKNOWLEDGMENTS

We thank Francesco Brunetti for developing the eMSQOL-29 software, and all MS patients who participated in the study.

DECLARATION OF CONFLICTING INTERESTS

None.

FUNDING

This work was supported by the Fondazione Italiana Sclerosi Multipla (FISM) grant number 2013/R/20.

REFERENCES

1 Rothwell PM, McDowell Z, Wong CK, et al. Doctors and patients don't agree: Cross sectional study of patients' and doctors' perceptions and assessments of disability in multiple sclerosis. *BMJ* 1997; 314: 1580-1583.

2 Solari A. Role of health-related quality of life measures in the routine care of people with multiple sclerosis. *Health Qual Life Outcomes* 2005; 3: 16. DOI: 10.1186/1477-7525-3-16.

3 Khurana V, Sharma H, Afroz N, et al. Patient-reported outcomes in multiple sclerosis: a systematic comparison of available measures. *Eur J Neurol* 2017; 24: 1099-1107.

4 Skevington SM, Lotfy M, O'Connell KA. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. *Qual Life Res* 2004; 13: 299-310.

5 Sloan JA, Aaronson N, Cappelleri JC, et al. Assessing the clinical significance of single items relative to summated scores*. Mayo Clin Proc* 2002; 77: 479-487.

6 Gold SM, Heesen C, Schulz H, et al. Disease specific quality of life instruments in multiple sclerosis: validation of the Hamburg Quality of Life Questionnaire in Multiple Sclerosis (HAQUAMS*). Mult Scler* 2001; 7: 119-130.

7 Ford HL, Gerry E, Tennant A, et al. Developing a disease-specific quality of life measure for people with multiple sclerosis. *Clin Rehabil* 2001; 15: 247–258.

8 Marrie RA, Miller DM, Chelune GJ, et al. Validity and reliability of the MSQLI in cognitively impaired patients with multiple sclerosis. *Mult Scler* 2003; 9: 621-626.

9 Simeoni M, Auquier P, Fernandez O; MusiQol study group. Validation of the Multiple Sclerosis International Quality of Life questionnaire. *Mult Scler* 2008; 14: 219-230.

10 Vickrey BG, Hays RD, Harooni R, et al. A health-related quality of life measure for multiple sclerosis. *Qual Life Res* 1995; 4: 187-206.

11 Rosato R, Testa S, Bertolotto A, et al. Development of a short version of MSQOL-54 using factor analysis and item response theory. *PLoS One* 2016; 11: e0153466. DOI: 10.1371/journal.pone.0153466.

12 Hays R, Stewart A. The structure of self-reported health in chronic disease patients. *Psychological Assessment* 1990; 2: 22-30.

13 Solari A, Filippini G, Mendozzi L, et al. Validation of Italian multiple sclerosis quality of life 54 questionnaire. *J Neurol Neurosurg Psychiatry* 1999; 67: 158-162.

14 Coons SJ, Gwaltney CJ, Hays RD, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. *Value Health* 2009; 12: 419-429.

15 Muehlhausen W, Doll H, Quadri N, et al. Equivalence of electronic and paper administration of patient-reported outcome measures: a systematic review and meta-analysis of studies conducted between 2007 and 2013. *Health Qual Life Outcomes* 2015; 13: 167. DOI: 10.1186/s12955-015-0362-x.

16 Eremenco S, Coons SJ, Paty J, et al. PRO data collection in clinical trials using mixed modes: report of the ISPOR PRO mixed modes good research practices task force. *Value Health* 2014; 17: 501-516.

17 Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann Neurol* 2011; 69: 292–302.

18 Cella DF, Dineen K, Arnason B, et al. Validation of the Functional Assessment of Multiple Sclerosis quality of life instrument. *Neurology* 1996; 47: 129-139.

19 Johnson JA, Coons SJ, Ergo A, et al. Valuation of EuroQOL (EQ-5D) health states in an adult US sample. *Pharmacoeconomics* 1998; 13: 421-433.

20 König HH, Bernert S, Angermeyer MC, et al. Comparison of population health status in six european countries: results of a representative survey using the EQ-5D questionnaire. *Med Care* 2009; 47: 255-261.

21 Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand* 1983; 67: 361–370.

22 Costantini M, Musso M, Viterbori P, et al. Detecting psychological distress in cancer patients: validity of the Italian version of the Hospital Anxiety and Depression Scale. *Support Care Cancer* 1999; 7: 121-127.

23 Provinciali L, Ceravolo MG, Bartolini M, et al. A multidimensional assessment of multiple sclerosis: relationships between disability domains. *Acta Neurol Scand* 1999; 100: 156-162.

24 Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 1983; 33: 1444-1452.

25 Drake AS, Weinstock-Guttman D, Morrow SA, et al. Psychometris and normative data for the Multiple Sclerosis Functional Composite: Replacing the PASAT with the Symbol Digit Modalities Test. *Mult Scler* 2010: 16: 228-237.

26 Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; 86: 420-428.

27 Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med* 1998; 17: 101-110.

28 Kadel RP, Kip KE. A SAS Macro to Compute Effect Size (Cohen's) and its Confidence Interval from Raw Survey Data. Proceedings of the annual Southeastern SAS Users Group Conference, 2012. http://analytics.ncsu.edu/sesug/2012/SD-06.pdf (accessed 15 December 2017).

29 Mitchell AJ, Benito-Leon J, Gonzalez JM, Rivera-Navarro J, et al. Quality of life and its assessment in multiple sclerosis: integrating physical and psychological components of wellbeing. *Lancet Neurol* 2005; 4: 556-566.

30 Moore F, Vickrey B, Fortin K, Lee L. Two multiple sclerosis quality-of-life measures: Comparison in a national sample. *Can J Neurol Sci* 2015; 42: 55-63.

31 Spilker B, Revicki DA. Taxonomy of quality of life. In: Spilker B (ed.), Quality of Life and Pharmacoeconomics in Clinical Trials. Philadelphia: Lippincott-Raven Publishers, 1996: 25–31.

32 Foley FW, Zemon V, Campagnolo D, et al. The Multiple Sclerosis Intimacy and Sexuality Questionnaire -- re-validation and development of a 15-item version with a large US sample. *Mult Scler* 2013; 19: 1197-1203.

33. Middleton LS, Denney DR, Lynch SG, Parmenter B. The relationship between perceived and objective cognitive functioning in multiple sclerosis. Arch Clin Neuropsychol. 2006 Aug;21(5):487-94.

34. Gold SM, Schulz H, Monch A, Schulz KH, Heesen C. Cognitive impairment in multiple sclerosis does not affect reliability and validity of self-report health measures. Multiple Sclerosis 2003; 9:

35. Bennett AV, Dueck AC, Mitchell SA, et al. Mode equivalence and acceptability of tablet computer-, interactive voice response system-, and paper-based administration of the U.S. National Cancer Institute's Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). *Health Qual Life Outcomes* 2016; 14: 24.

36. Guyatt GH, Osoba D, Wu AW, et al. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc* 2002; 77: 371–383.

37. Schwartz CE, Sprangers MA. Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Soc Sci Med* 1999; 48: 1531–1548.

38. Lavallee DC, Chenok KE, Love RM, et al. Incorporating patient-reported outcomes into health care to engage patients and enhance care. Health Affairs 2016; 35: 575-582.

Table 1. Characteristics of participants (N=623) by centre.

Characteristic	Milan (n=100)	Orbassano (n=147)	Catania (n=166)	Rome (n=110)	Chieti (n=100)	Total (n=623)
Women[1]	66 (66.0)	94 (63.9)	120 (72.3)	78 (70.9)	66 (66.0)	424 (68.1)
Age (years)[2]	39.2, 9.4 (22-59)	46.9, 12.2 (21-77)	42.0, 10.6 (20-69)	51.8, 11.2 (28-78)	40.9, 9.0 (22-61)	44.3, 11.5 (20-78)
Education[1]: Primary	21 (21.0)	36 (24.5)	29 (17.5)	15 (13.6)	19 (19.0)	120 (19.3)
Secondary	41 (41.0)	75 (51.0)	93 (56.0)	44 (40.0)	53 (53.0)	306 (49.1)
College/University	38 (38.0)	36 (24.5)	44 (26.5)	51 (46.4)	28 (28.0)	197 (31.6)
Married (vs. not-married)	55 (55.0)	95 (64.6)	116 (69.9)	65 (59.1)	61 (61.0)	392 (62.9)
Years from MS diagnosis[3,*]	7 (1-31)	11 (1-47)	7 (1-32)	18 (1-42)	7 (1-26)	10 (1-47)
Disease course[&]: RR	90 (95.7)	120 (81.6)	143 (87.2)	58 (52.7)	87 (14.4)	498 (82.2)
PP	1 (1.1)	11 (7.5)	7 (4.3)	7 (6.4)	3 (3.3)	29 (4.8)
SP	3 (3.2)	16 (10.9)	14 (8.5)	45 (40.9)	1 (1.1)	79 (13.0)
EDSS score[3,*]	1.5 (0.0-8.0)	2.0 (0.0-7.5)	2.0 (0.0-7.0)	6.0 (1.0-9.0)	2.5 (0.0-6.5)	2.5 (0.0-9.0)
SDMT score[2,*]	51.0, 14.2 (17.0-93.0)	44.5, 14.7 (0.0-84.0)	50.3, 11.2 (19.0-71.0)	32.5, 15.1 (1.0-68.0)	56.4, 10.7 (28.0-81.0)	46.7, 15.3 (0.0-93.0)
HADS-Anxiety score[2,^]	7.0, 4.1 (0.0-17.0)	6.5, 4.1 (0.0-19.0)	4.5, 3.4 (0.0-11.0)	6.0, 4.2 (0.0-18.0)	6.7, 3.8 (0.0-14.0)	6.0, 4.0 (0.0-19.0)
HADS-Depression score[2,^]	4.2, 3.8 (0.0-19.0)	3.9, 3.3 (0.0-15.0)	3.8, 3.8 (0.0-13.0)	4.1, 3.8 (0.0-16.0)	5.4, 4.2 (0.0-18.0)	4.2, 3.8 (0.0-19.0)
FAMS total score[2,°]	131.3, 29.3 (37-176)	126.8, 25.4 (49-169)	126.5, 26.2 (49-169)	113.5, 27.0 (45-163)	128.0, 24.2 (77-174)	125.3, 26.9 (37-176)
EQ-5D-3L score[2,°]	0.9, 0.1 (0.0-1.00)	0.8, 0.2 (-0.1-1.00)	0.9, 0.1 (0.7-1.00)	0.7, 0.3 (-0.4-1.00)	0.9, 0.1 (0.3-1.00)	0.8, 0.2 (-0.4-1.00)
EQ-VAS[2]	73.6, 16.6 (30-100)	71.0, 18.8 (0-100)	69.7, 16.3 (30-100)	61.5, 22.7 (0-100)	74.2, 14.3 (37-100)	69.9, 18.4 (0-100)
Full-set of sexual items[1]	55 (55.0)	76 (51.7)	116 (69.9)	19 (17.3)	65 (65.0)	331 (53.1)

[1]Number (%); [2]Mean, standard deviation (min-max); [3]Median (min-max); *Valid cases N=606; ^Valid cases N=614; °Valid cases N=620; [&]Valid cases N=606.

EDSS, Expanded Disability Status Scale; EQ-5D-3L, European Quality of life Five Dimensions-3L; MS, FAMS; Functional Assessment of Multiple Sclerosis; HADS, Hospital Anxiety Depression Scale; Multiple Sclerosis; RR, Relapsing Remitting; PP, Primary Progressive; SDMT, Symbol Digit Modality Test; SP, Secondary progressive; VAS, Visual Analogue Scale

Table 2. Patient characteristics by response to the filter question on sexual activity.

Characteristic	Full-set of sexual items (N=309)	Single sexual item (N=273)	Total (N=582)
Women[1]	214 (69.3)	182 (66.7)	396 (68.0)
Age (years)[2]	41.1, 9.9 (21-78)	47.7, 12.4 (21-77)	44.2, 11.6 (21-78)
Education[1]			
Primary	47 (15.2)	68 (24.9)	115 (19.8)
Secondary	164 (53.1)	120 (44.0)	284 (48.8)
College/University	98 (31.7)	85 (31.1)	183 (31.4)
Married (vs. not-married)	209 (67.6)	152 (55.7)	361 (62.0)
Years from MS diagnosis[3,*]	7 (1–34)	12 (1-47)	10 (1–47)
EDSS score[3,*]	2 (0-7.5)	3.5 (0-9)	2 (0–9)
SDMT score[2,*]	51.4, 12.8 (0–93)	41.1, 16.2 (0–83)	46.6, 15.3 (0–93)
HADS Anxiety score[2,^]	5.2, 3.8 (0–17)	6.9, 4.1 (0–19)	6.0, 4.0 (0–19)
HADS Depression score[2,^]	3.3, 3.3 (0–14)	5.0, 3.8 (0–16)	4.1, 3.7 (0–16)
FAMS total score[2,°]	134.3, 23.0 (56–176)	115.9, 27.4 (45–174)	125.7, 26.7 (45–176)
EQ-5D-3L score[2,°]	0.9, 0.1 (-0.1–1.0)	0.8, 0.2 (-0.4–1.0)	0.8, 0.2 (-0.04–1.0)
MS Centre (%): Milan	52 (16.8)	40 (14.7)	92 (15.8)
Orbassano	75 (24.3)	66 (24.2)	141 (24.2)
Catania	111 (35.9)	48 (17.6)	159 (27.3)
Rome	18 (5.8)	89 (32.6)	107 (18.4)
Chieti	53 (17.2)	30 (11.0)	83 (14.3)

[1] Number (%); [2] Mean, standard deviation (min-max); [3] Median (min-max)

EDSS, Expanded Disability Status Scale; EQ-5D-3L, European Quality of life Five Dimensions-3L FAMS, Functional Assessment of Multiple Sclerosis; HADS, Hospital Anxiety Depression Scale; MS, Multiple sclerosis; SDMT, Symbol Digit Modality Scale.

Table 3. Summary statistics of the MSQOL-29 subscale scores.

Subscale	Valid cases	Mean (SD)	Q1	Q3	Cronbach's alpha	Percentage of patients scoring min/max
Physical function	582	66.8 (35.7)	41.7	100	0.89	9.8/36.8
Bodily pain	582	76.6 (25.1)	60.0	100	0.89	0.5/36.9
Emotional wellbeing	581	65.0 (18.8)	53.3	76.7	0.88	0.2/ 2.8
Energy	581	52.0 (20.6)	41.7	66.7	0.88	1.2/ 1.0
Cognitive function	582	67.8 (22.7)	53.3	85.0	0.89	0.2/12.9
Health distress	581	74.5 (23.0)	60.0	93.3	0.88	1.2/20.3
Sexual function	581	79.5 (29.6)	66.7	100	0.89	5.5/54.7
Sexual function[§]	309	88.0 (19.6)	83.4	100	0.90	0.3/59.2
Health perceptions	581	48.8 (30.7)	33.3	66.7	-	18.9/10.8
Social function	579	63.5 (27.4)	50.0	75.0	-	4.2/22.1
Overall quality of life	549	66.1 (17.3)	50.0	80.0	-	0.7/ 0.9
Change in health	582	52.0 (23.7)	50.0	75.0	-	4.5/ 9.3

* Single item

[§] Values of patients who completed the full-set of the 'sexual function' subscale.

SD, standard deviation; Q1, lower quartile; Q3, upper quartile.

Table 4. MSQOL-29 subscale loadings and weights for the Mental (MHC) and Physical (PHC) Health Composite scores obtained from confirmatory factor analysis.

Subscale	Standardized regression coefficient (loading)		Weight	
	MHC	PHC	MHC	PHC
Bodily pain	0.61		0.14	
Emotional wellbeing	0.73		0.17	
Cognitive function	0.60		0.14	
Social function (single item)	0.74		0.17	
Energy	0.83		0.19	
Health distress	0.77		0.18	
Physical function		0.70		0.22
Sexual function		0.52		0.17
Health perceptions (single item)		0.62		0.20
Overall quality of life (single item)		0.78		0.25
Change in health (single item)		0.51		0.16

Table 5. Pearson's correlation coefficients of MSQOL-29 and FAMS, HADS, EQ-5D-3L and SDMT scales/subscales. Values for scales/subscales addressing similar constructs are reported in bold. All correlations were statistically significant at p <0.001.

MSQOL-29	FAMS						HADS		EQ-5D		SDMT
Subscale	MO	SY	EWB	GC	TF	FSWB	A	D	Index	VAS	
Physical function	**0.86**	0.43	0.58	0.51	0.55	0.24	-0.26	-0.45	0.62	0.60	0.56
Bodily pain	0.48	**0.66**	0.41	0.31	0.44	0.20	-0.40	-0.36	0.44	0.42	0.22
Emotional wellbeing	0.42	0.37	**0.62**	0.58	0.45	0.39	**-0.59**	**-0.56**	0.36	0.43	0.19
Energy	0.61	0.52	**0.58**	0.53	**0.62**	0.29	-0.48	-0.52	0.50	0.53	0.28
Cognitive function	0.40	0.42	0.42	0.45	**0.75**	0.35	-0.41	-0.51	0.26	0.40	**0.25**
Health distress	0.57	0.46	**0.69**	0.59	0.50	0.36	-0.56	-0.53	0.43	0.47	0.29
Sexual function	0.41	0.28	0.43	0.37	0.44	0.23	-0.35	-0.40	0.31	0.31	0.25
Health perceptions	0.48	0.35	0.48	0.39	0.43	0.19	-0.35	-0.36	0.38	0.41	0.23
Social function	0.57	0.38	0.55	0.51	0.52	**0.38**	-0.40	-0.48	0.42	0.45	0.33
Overall quality of life	0.61	0.45	0.63	**0.65**	0.59	0.44	-0.47	-0.61	**0.48**	**0.66**	0.36
Change in health	0.44	0.36	0.39	0.33	0.40	0.16	-0.21	-0.31	0.33	0.45	0.24

EWB, emotional well-being; EQ-5D, European Quality of life Five Dimensions; FAMS, Functional Assessment of Multiple Sclerosis; FSWB, family social well-being; GC, general contentment; HADS, Hospital Anxiety and Depression Scale; MO, mobility; SDMT, Symbol Digit Modalities Test; SY, symptoms; TF, thinking/fatigue; VAS, visual analogue scale

Table 6. Distribution of eMSQOL-29 subscale scores, and measures to assess equivalence. CI is confidence interval; ICC is intraclass correlation coefficient; SD is standard deviation.

	N	Mean (SD) eMSQOL-29	Mean (95% CI) difference paper-eMSQOL-29	Effect size (95% CI) paper - eMSQOL-29	ICC (95% CI)
Multi-item subscales:					
Physical function	220	74.4 (30.1)	-0.08 (-1.77 – 1.62)	-0.006 (-0.063 – 0.051)	0.95 (0.94 – 0.96)
Bodily pain	220	75.7 (24.3)	0.29 (-1.90 – 2.48)	0.018 (-0.072 – 0.107)	0.87 (0.83 – 0.90)
Emotional wellbeing	219	64.1 (19.3)	0.96 (-0.85 – 2.77)	0.071 (-0.024 – 0.165)	0.85 (0.81 – 0.89)
Energy	219	51.5 (21.0)	-0.04 -(1.79 – 1.72)	-0.003 (-0.087 – 0.081)	0.89 (0.85 – 0.91)
Cognitive function	220	68.0 (22.4)	1.23 (-0.71 – 3.18)	0.084 (-0.001 – 0.170)	0.88 (0.85 – 0.91)
Health distress	219	74.5 (21.2)	0.84 (-1.32 – 3.00)	0.052 (-0.044 – 0.148)	0.85 (0.80 – 0.88)
Sexual function	219	78.2 (30.0)	1.58 (-1.48 – 4.65)	0.069 (-0.031 – 0.169)	0.84 (0.79 – 0.87)
Single-item subscales:					
Health perceptions	219	47.9 (28.0)	1.52 (-1.79 – 4.84)	0.061 (-0.055 – 0.177)	0.62 (0.54 – 0.70)
Social function	219	65.3 (27.1)	1.94 (-1.62 – 5.50)	0.073 (-0.057 – 0.202)	0.52 (0.42 – 0.61)
Overall quality of life	208	68.9 (17.0)	1.35 (-0.25 – 2.94)	0.115 (0.022 – 0.208)	0.77 (0.71 – 0.82)
Change in health	215	54.8 (23.8)	0.58 (-2.20 – 3.37)	0.028 (-0.086 – 0.142)	0.63 (0.55 – 0.71)

Table 7. Repeated-measure linear regression models of the 11 eMSQOL-29 subscale scores (dependent variables). Each model includes the following independent variables: version (paper vs. electronic), order (first vs. second administration), sequence (version per order), centre (Milan vs. Orbassano), sex, age, Expanded Disability Status Scale score (≤2.5 vs. >2.5), and disease course (relapsing-remitting vs. primary or secondary progressive).

Subscale	Version		Order of administration		Sequence	
	β (SE)	P value	β (SE)	P value	β (SE)	P value
Physical function	-0.01 (0.9)	0.99	0.08(0.9)	0.92	12.50 (10.6)	0.24
Bodily pain	0.21 (1.1)	0.85	1.47 (1.1)	0.19	5.23 (2.9)	0.07
Emotional wellbeing	0.98 (0.9)	0.29	-1.05 (0.9)	0.25	1.26 (2.4)	0.60
Energy	0.08 (0.9)	0.92	-1.69 (0.9)	0.06	1.87 (2.5)	0.46
Cognitive function	1.34 (1.0)	0.17	-1.57 (1.0)	0.11	4.37 (2.8)	0.12
Health distress	0.66 (1.1)	0.55	0.19 (1.1)	0.86	3.43 (2.6)	0.19
Sexual function	1.70 (1.5)	0.27	-2.77 (1.5)	0.07	-0.55 (3.5)	0.88
Health perceptions	1.43 (1.7)	0.39	1.73 (1.7)	0.30	-0.56 (3.3)	0.86
Social function	1.81 (1.8)	0.32	0.66 (1. 8)	0.71	3.59 (3.1)	0.24
Overall quality of life	1.47 (0.8)	0.07	0.13 (0.8)	0.87	1.66 (2.0)	0.41
Change in health	0.65 (1.4)	0.65	0.48 (1.4)	0.73	1.65 (2.7)	0.54

All estimates are adjusted for centre (Milan, Orbassano), sex, age, Expanded Disability Status Scale and diagnosis.
SE, standard error.